



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :  C12Q 1/68		A1	(11) International Publication Number: WO 00/39332  (43) International Publication Date: 6 July 2000 (06.07.00)
<p>(21) International Application Number: PCT/GB99/04380</p> <p>(22) International Filing Date: 22 December 1999 (22.12.99)</p> <p>(30) Priority Data: 9828619.8 23 December 1998 (23.12.98) GB</p> <p>(71) Applicant (for all designated States except US): JANSSEN PHARMACEUTICA N.V. [BE/BE]; Turnhoutseweg 30, B-2340 Beerse (BE).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (for US only): PAULUSSEN, Aimee, Dymphne, Catherine [NL/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B-2340 Beerse (BE). ARMSTRONG, Martin [GB/GB]; Pippins, Grove Road, Wickambreux, Canterbury, Kent CT3 1SJ (GB).</p> <p>(74) Agent: BOULT WADE TENNANT; Verulam Gardens, 70 Gray's Inn Road, London WC1X 8BT (GB).</p>		<p>(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FL, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</p> <p><b>Published</b> With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</p>	
<p>(54) Title: GENOTYPING CYTOCHROME EXPRESSION</p> <p>(57) Abstract</p> <p>There is disclosed a method of identifying subjects having a high or low drug metabolising phenotype associated with cytochrome CYP3A5 expression, which method comprises screening genomic DNA from said subject for the presence or absence of one or more polymorphic variants in a transcription regulatory region of the sequence encoding CYP3A5. Oligonucleotide molecules for carrying out the screening are also provided.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

## GENOTYPING CYTOCHROME EXPRESSION

The present invention is concerned with an assay and, in particular, with an assay for genotyping a 5 polymorphism predictive of a phenotype associated with cytochrome expression, in this case CYP3A5.

The cytochrome P450 subfamily CYP3A represents one of the most important families of the P450 superfamily 10 and plays a major role in the metabolism of an ever expanding list of therapeutic compounds (23, 24). This family comprises the most abundantly expressed P450s in human livers, and is responsible for the metabolism of over 50% of all clinically used drugs, including 15 the dihydropyridines, cyclosporin, erythromycin and barbiturates (1). Wide inter-individual variation in the metabolism of CYP3A substrates has been noted and is a factor in determining individual drug efficacy. Evidence also exists for the metabolism of an array of 20 lipophilic environmental pollutants, including the activation of pro-carcinogens such as aflatoxin B1 by members of this subfamily (2).

Presently, four CYP3A cDNAs have been identified in 25 humans, CYP3A3, CYP3A4, CYP3A5 and CYP3A7. It is believed that CYP3A3 represents an allelic variant of CYP3A4, whilst CYP3A4 and CYP3A7 are found only in 30 human adult and fetal livers respectively (3). Initial experiments suggested that a polymorphism existed in CYP3A4 (4). However other studies, whilst confirming a wide range of inter-individual variation in CYP3A4 expression have failed to confirm the original 35 bimodality (5, 6). Overlapping substrate specificities between CYP3A5 and CYP3A5 have previously made it difficult to separate metabolism by these isoforms;

- 3 -

metabolism driven towards the 1-OHM route and therefore show a higher ratio of 1-OHM/4-OHM than those containing only CYP3A4. The present inventors have now established that two polymorphisms, located in putative transcriptional regulatory regions, which caused increased CYP3A5 gene expression and metabolic activity are linked and have developed assays for their detection. These assays will allow prediction of inter-individual variability in response to drugs metabolised by this isoform, as well as facilitating disease association studies.

Therefore, according to a first aspect of the present invention there is provided a method of identifying subjects having a high or low drug metabolising phenotype associated with cytochrome CYP3A5 expression, which method comprises screening for the presence or absence in the genome of a subject a polymorphic variant in a transcription regulatory region, such as, a promoter or enhancer adjacent the region encoding CYP3A5. Preferably, the method involves screening for a variant in a recognition site for a transcription factor of said regulatory region, and even more preferably in an activator protein-3 motif or a basic transcription element. Even more preferably, the method involves screening for a variant at any one of positions -475 or -147 of the DNA of the 5' flanking region adjacent to the region encoding CYP3A5 the sequence of which flanking region is illustrated in Figure 7 and preferably, for both the variants at positions -475 and -147.

In one embodiment of the method of the invention genomic DNA is amplified, preferably by the polymerase

chain reaction using oligonucleotide molecules capable of hybridising selectively to the wild type sequence or the variant sequences, such that generation of amplified DNA from said molecules will indicate whether said wild type or mutation is present. In this method PCR primers hybridise either to the mutated or wild type sequence, but not both. Amplification of the DNA of the respective mutation or wild type genotype using the respective primers will provide an indication of the presence of the wild type or mutated nucleotide mutations.

A further method of the invention advantageously utilises oligonucleotide molecules as primers which, in addition to hybridising to the site of interest, are capable of introducing a restriction site which is absent in either the wild type sequence or polymorphic variants. Therefore, according to a further aspect of the invention, there is provided a method of identifying subjects having a high or low drug metabolising phenotype associated with CYP 3A5 expression, which method comprises 1) amplifying genomic DNA from a subject using oligonucleotide molecules capable of hybridising to the wild type sequence and/or to the polymorphic variant sequence at a location being analysed, which molecules are such that they can introduce a restriction site at said location which is not present in the wild type or variant sequences, and 2) subjecting amplified DNA from step 1 to a restriction enzyme which cleaves the DNA at said restriction site to provide a restriction digest indicative of the presence or absence of said variant.

The method preferably comprises amplifying DNA in a

recognition site for a transcription factor of said regulatory region and preferably in an activator protein-3 motif (AP-3) and/or basic transcription element (BTE). Preferably, the method comprises amplifying DNA spanning any of position -475 or -147, of the regulatory region of CYP 3A5, the sequence of which is illustrated in Figure 7.

The polymorphisms at the positions identified in each of the methods according to the invention comprise T<sub>-475</sub> - G and A<sub>-147</sub> - G. As presented in the Examples below, the molecule used to detect the variation at A<sub>-147</sub> - G is capable of introducing a restriction site for the enzyme *Tai* I only when the wild type A nucleotide is present at position -147. Alternatively, the molecule used to detect the T<sub>-475</sub> - G nucleotide variation is capable of introducing a restriction site for the enzyme *Alu* I only when the wild type T nucleotide is present at position -475.

In this embodiment an example of suitable primers is any of 3A5F1 GGGTCTGTCTGGCTGCGC and 3A5F2 (GGGGTCTGTCTGGCTGAGC) and 3A5R1 (TTTATGTGCTGGAGAAGGACG).

Using oligonucleotide mismatch primer 3A5R1 creates a *Tai* I recognition site only when the wild type A nucleotide is present at position -147. Digestion of the 369bp product with *Tai* I yields fragments of 349 and 20bp for the wild type sequence, whilst the product remains undigested if a mutant, such as the G nucleotide, is present (Figure 2). Similarly, for the detection of the T<sub>-475</sub>G mutation a second oligonucleotide mismatch primer 3AF2 may be used.

This primer introduces a recognition site for the restriction enzyme *Alu* I when the wild type T is present at position -475, digesting the product to yield fragments of 318, 33 and 18bp. This site is lost when the mutant G nucleotide is present, yielding digestion products of 336 and 33bp (Figure 3).

Known techniques for the scoring of single nucleotide polymorphisms (see review by Schafer, A. J. and Hawkins, J. R. in *Nature Biotechnology*, Vol 16, pp33-39 (1998) include mass spectrometry, particularly matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS, see Roskey, M. T. et.al., 1996, *PNAS USA*, 93: 4724-4729), single nucleotide primer extension (Shumaker, J. M. et.al., 1996, *Hum. Mutat.*, 7: 346-354; Pastinen, T. et.al., 1997, *Genome Res.*, 7: 606-614) and DNA chips or microarrays (Underhill, P. A. et.al., 1996, *PNAS USA*, 93: 196-200; Gilles, P. N. et.al. *Nat. Biotech.*, 1999, 17: 365-370). The use of DNA chips or microarrays could enable simultaneous genotyping at many different polymorphic loci in a single individual or the simultaneous genotyping of a single polymorphic locus in multiple individuals.

In addition to the above, SNPs are commonly scored using PCR-SSCP based techniques, such as PCR-SSP using allele-specific primers (described by Bunce, 1995). If the SNP results in the abolition or creation of a restriction site then genotyping can be carried out by performing PCR using non-allele specific primers spanning the polymorphic site and digesting the resultant PCR product using the appropriate restriction enzyme. The known techniques for scoring polymorphisms are of general applicability and it would therefore be readily apparent to persons

skilled in the art that the known techniques could be adapted for the scoring of single nucleotide polymorphisms in the the regulatory region of CYP 3A5.

As would be readily apparent to those skilled in the art, genotyping is generally carried out on genomic DNA prepared from a suitable tissue sample obtained from the subject under test. Most preferably, genomic DNA is prepared from a blood sample, according to standard procedures which are well known in the art

Also provided by the present invention is an oligonucleotide of at least 10 contiguous nucleotides to detect polymorphic variants in a 5' regulatory region adjacent the sequence encoding cytochrome CYP3A5 associated with a high or low drug metabolising phenotype. The oligonucleotide is capable of hybridising to a region incorporating either a mutated or wild type nucleotide at position -475 or -147 of said flanking region, such that amplification of said positions will or will not proceed from said primer according to whether or not a polymorphic variant occurs at any of said positions.

The oligonucleotide molecules of the invention are preferably from 10 to 50 nucleotides in length, even more preferably from 20-30 nucleotides in length, and may be DNA, RNA or a synthetic nucleic acid, and may be chemically or biochemically modified or may contain non-natural or derivatized nucleotide bases, as will be readily appreciated by those skilled in the art. Possible modifications include, for example, the addition of isotopic or non-isotopic labels, substitution of one or more of the naturally occurring nucleotide bases with an analog, internucleotide

modifications such as uncharged linkages (e.g. methyl phosphonates, phosphoamidates, carbamates, etc.) or charged linkages (e.g. phosphorothioates, phosphorodithioates, etc.). Also included are synthetic molecules that mimic polynucleotides in their ability to bind to a designated sequence to form a stable hybrid. Such molecules are known in the art and include, for example, so-called peptide nucleic acids (PNAs) in which peptide linkages substitute for phosphate linkages in the backbone of the molecule. An oligonucleotide molecule according to the invention may be produced according to techniques well known in the art, such as by chemical synthesis or recombinant means.

The oligonucleotide molecules of the invention may be double stranded or single stranded but are preferably single stranded, in which case they may correspond to the sense strand or the antisense strand of the 5' regulatory region of CYP3A5. The oligonucleotides may advantageously be used as probes or as primers to initiate DNA synthesis/DNA amplification. They may also be used in diagnostic kits or the like for detecting the presence of one or more variants alleles of the regulatory region of CYP3A5. These tests generally comprise contacting the probe with a sample of test nucleic acid (usually genomic DNA) under hybridising conditions and detecting for the presence of any duplex or triplex formation between the probe and complementary nucleic acid in the sample. The probes may be anchored to a solid support to facilitate their use in the detection of these variants. Preferably, they are present on an array so that multiple probes can simultaneously hybridize to a single sample of target nucleic acid. The probes can be spotted onto the array or

- 9 -

synthesised *in situ* on the array. (See Lockhart et al., *Nature Biotechnology*, vol. 14, December 1996 "Expression monitoring by hybridisation to high density oligonucleotide arrays". A single array can contain more than 100, 500 or even 1,000 different probes in discrete locations. Preferably, the oligonucleotides comprise any of the primers 3A5F1, 3A5F2 and 3A5R1 as defined herein.

Also provided is a kit to perform the method according to the invention. Preferably, the kit will comprise an oligonucleotide as described herein and even more preferably the kit will further comprise one or more restriction enzymes capable of distinguishing between wild-type or polymorphic variants as defined herein. Preferably, the restriction enzyme comprises *Tai* I or *Alu* I.

According to a further aspect of the invention there is also provided a method of identifying toxic or mutagenic effects of a test compound, such as, a drug, toxin or procarcinogen metabolised by CYP3A5 the method comprising contacting each of a cell having a high drug metabolising phenotype and a cell having a low metabolising phenotype associated with cytochrome CYP3A5 expression, with said test compound and identifying the effects of said compound on each of said high or low drug metabolising phenotype cells or other cells sensitive to said compound. An even further aspect comprises a method of diagnosing susceptibility of an individual to a disease associated with environmental toxins or procarcinogens metabolised by CYP3A5, the method comprising the steps of 1) providing a sample containing DNA, and 2) identifying the presence or absence of a mutation in a

- 10 -

transcription regulatory region adjacent to the DNA sequence encoding CYP3A5 using a reagent capable of distinguishing the presence or absence of a nucleotide in said regulatory site. According to this aspect of the invention, the mutation occurs in a recognition site for a transcription factor of said regulatory region and preferably in an activator protein-3 motif (AP-3) and/or a basic transcription element (BTE). Preferably, the mutation occurs at any of positions -475 and -147 of the regulatory region and even more preferably at both positions where the mutation may be T<sub>-475</sub>G or A<sub>-147</sub>G.

Advantageously, it is also envisaged that the regulatory region of the 5' flanking region can be used to identify or purify transcription factors which bind to the 5' region including the respective polymorphic variants. Thus, according to a further aspect of the invention, there is provided a method of identifying transcription factors capable of binding to a DNA fragment from a transcription regulatory region adjacent DNA encoding cytochrome CYP3A5, said method comprising contacting said DNA fragment including said transcription regulatory region with potential transcription factors and identifying any transcription factor complexed to said DNA fragments.

Using the transcription regulatory fragment it is possible to identify compounds or agents which exhibit or exert their effect on the transcription regulatory region of CYP3A5. Thus, there is provided according to this aspect of the invention a method of identifying compounds acting on a transcription regulatory region adjacent to a DNA sequence encoding CYP3A5, the method comprising transforming a cell with

- 11 -

a DNA construct comprising the sequence of said regulatory region, and which regulatory region is operably linked to a sequence encoding a reporter molecule, contacting said cell with a test compound and identifying any expression of said reporter molecule. Preferably, said cell is expressing CYP3A5 or is showing CYP3A5 activity.

Also provided by the invention is a method of purification of transcription factors from a sample which are capable of binding to DNA from a transcription regulatory region adjacent a DNA sequence encoding cytochrome CYP3A5, the method comprising contacting a DNA fragment including said transcriptional regulatory region with a mixture of transcription factors and identifying any complexes of said transcription factors and said fragment.

An even further aspect of the invention comprises a method of providing a measure of activity of a transcription regulatory region adjacent to DNA encoding cytochrome CYP3A5 or alternatively a method of identifying a mutation which alters the activity of the transcription regulatory region the method comprising providing a DNA construct having a sequence encoding a reporter molecule operably linked to a DNA fragment comprising said regulatory region, and introducing said construct into a cell and monitoring for the level of expression of said reporter molecule. When the method is used to identify a variant which alters the activity of the transcription regulatory control region, the method may include the further step of comparing the levels of expression of a wild type and a polymorphic regulatory region as described herein.

According to each of the aspects of the invention, the regulatory region includes a polymorphic variation, preferably in a recognition site for a transcription factor of said regulatory region, and preferably in an activator protein-3 motif (AP-3) and/or a basic transcription element (BTE). In a preferred embodiment the variant occurs at position -475 or -147 of the region flanking the sequence encoding CYP3A5, and which region is illustrated in Figure 7. Preferably, both the variants are present.

*CRP  
Heit*

---

The methods of the present invention will be particularly valuable to establish, prior to treatment with a drug, whether the drug will be effectively metabolised by the patient.

---

The invention may be more clearly understood by the following example with reference to the accompanying drawings wherein

20

**Fig. 1a:** is an illustration of the relationship between midazolam metabolic ratio and genotype for the linked A<sub>-147</sub>G and T<sub>-475</sub>G mutations in the 5' flanking region of the CYP3A5 gene. Midazolam metabolic ratio = 1-0HM/4-OHM, wt = samples with the wild type sequence in the 5' flanking region as previously published (11), Het = samples heterozygous for the linked polymorphisms, A<sub>-147</sub>G and T<sub>-475</sub>G.

25

**Fig. 1b:** is an illustration of the relationship between CYP3A5 mRNA expression and the linked A<sub>-147</sub>G and T<sub>-475</sub>G mutations in the 5' flanking region of CYP3A5. Relative Ct

30

35

CLAIMS

1. A method of identifying subjects having a high or low drug metabolising phenotype associated with cytochrome CYP3A5 expression, which method comprises the steps of:

screening genomic DNA from said subject for the presence or absence of one or more polymorphic variants in a transcription regulatory region of the sequence encoding CYP3A5 characteristic of a high drug metabolising phenotype.

2. A method of screening human subjects for suitability for treatment with a drug metabolised by CYP3A5 comprising screening for the presence or absence of one or more polymorphic variants in a transcription regulatory region of the sequence encoding CYP3A5 characteristic of a high drug metabolising phenotype.

3. A method according to claim 1 or 2 comprising screening for said one or more variants in a recognition site for a transcription factor of said regulatory region.

4. A method according to any of claims 1 to 3 comprising screening for said one or more variants in an activator protein-3 motif (AP-3) and/or basic transcription element (BTE).

5. A method according to any of claims 1 to 4, comprising screening for said one or more variants at any one of positions -475 or -147 of the transcription regulatory region of the sequence encoding CYP3A5 the sequence of which region is illustrated in Figure 7.

6. A method according to claim 5 comprising screening for both of said variants at position -475 or -147 of said transcriptional regulatory region of CYP3A5.

7. A method according to any of claims 1 to 5 wherein said DNA is amplified using oligonucleotide molecules which are capable of hybridising selectively to the wild type or variant sequences respectively such that generation of amplified DNA from said respective molecules will indicate whether said wild type or said variant is present.

8. A method of identifying one or more polymorphic variants in a transcription regulatory region of DNA encoding cytochrome CYP3A5 said method comprising the steps of:

- 1) subjecting the sample DNA to amplification using oligonucleotide molecules which are capable of selectively hybridising to the wild type and/or said one or more variant sequences, which molecules are such that they can introduce a restriction site in one of said amplified wild type or variant sequences, and
- 2) subjecting amplified DNA from step 1 to restriction with an enzyme which cleaves at said restriction site to provide a restriction digest indicative of the presence or absence of said mutation.

9. A method according to claim 8 wherein said molecule introduces a restriction site in a region corresponding to a recognition site for a transcription factor of said regulatory region.

10. A method according to claim 8 or 9 wherein said molecule introduces a restriction site in a region corresponding to an activator protein-3 motif (AP-3) and/or a basic transcription element (BTE).

11. A method according to claim 10 wherein said molecule is capable of introducing a restriction site only when the wild type A nucleotide is present at position -147 of the transcription regulatory region.

12. A method according to claim 11 wherein said restriction site is for the Tai I restriction enzyme.

13. A method according to claim 11 or 12 wherein said oligonucleotide molecule comprises the sequence designated 3A5R1 illustrated in Figure 6.

14. A method according to claim 10 wherein said molecule is capable of introducing a restriction site when the wild type T nucleotide is present at position -475 of the regulatory control region.

15. A method according to claim 14 wherein said restriction site is for the Alu I enzyme.

16. A method according to claim 14 or 15 wherein said molecule comprises the sequence designated 3A5F2 illustrated in Figure 6.

17. An oligonucleotide molecule of at least 10 contiguous nucleotides for use in amplification of a DNA sequence to detect a wild type or polymorphic variant in a transcription regulatory region of the sequence encoding cytochrome CYP3A5 said associated with a high or low drug metabolising phenotype

respectively, which molecule is capable of hybridising to a region incorporating either a polymorphic variant or wild type nucleotide in said region, such that amplification of said wild type and polymorphic variants will proceed from said molecule only when an oligonucleotide includes a sequence corresponding to either said wild type or polymorphic variant characteristic of a high drug metabolising phenotype.

18. A molecule according to claim 17 which is capable of hybridising to a recognition site for a transcription factor of said regulatory region.

19. A molecule according to claim 17 or 18 which is capable of hybridising to an activator protein-3 motif (AP-3) or a basic transcription element.

20. A molecule according to any of claims 17 to 19 which is capable of hybridising to a region comprising a polymorphic variant at any of positions -475 or -147 of the transcription regulatory region of the sequence encoding CYP3A5 illustrated in Figure 7.

21. A molecule according to any of claims 17 to 20 which comprises any of the sequences designated 3A5F1, 3A5F2 or 3A5R1 illustrated in Figure 6.

22. A kit for performing the method of any of claims 1 to 7 comprising an oligonucleotide molecule according to any of claims 17 to 21 and means for contacting said molecule and said transcription regulatory region of the sequence encoding CYP3A5.

23. A kit according to claim 22 further comprising a restriction enzyme capable of producing a

restriction digest for distinguishing between said variant or wild type sequences.

24. A kit according to claim 23 wherein said enzyme comprises any of *Tai I* or *Alu I*.

25. A method of identifying toxic or mutagenic effects of a test compound, such as, a drug, toxin or procarcinogen metabolised by CYP3A5 the method comprising contacting each of a cell having a high drug metabolising phenotype and a cell having a low metabolising phenotype associated with cytochrome CYP3A5 expression, with said test compound and identifying the effects of said compound on each of said high or low drug metabolising phenotype cells or other cells sensitive to said compound.

26. A method of diagnosing susceptibility of an individual to a disease associated with environmental toxins or procarcinogens metabolised by CYP3A5, which method comprises screening for the presence or absence of a polymorphic variant in a transcription regulatory region of the sequence encoding CYP3A5.

27. A method according to claim 26 comprising screening for said variant in a recognition site for a transcription factor of said regulatory region.

28. A method according to claim 26 or 27 comprising screening for said variant in an activator protein-3 motif (AP-3) and/or a basic transcription element (BTE) of said transcription regulatory region.

29. A method according to any of claims 26 to 28, comprising screening for said variant at any one

of position -475 or -147 of the transcription regulatory region of the sequence encoding CYP3A5 the sequence of which region is illustrated in Figure 7.

30. A method according to any of claims 26 to 29 comprising screening for both variants at position -475 or -147.

31. A method according to any of claims 26 to 30 comprising screening for the presence or absence of variants T<sub>-475</sub>G and A<sub>-147</sub>G in said transcriptional regulatory control region.

32. A method of providing a measure of activity of a transcription regulatory region of a DNA sequence encoding cytochrome CYP3A5 or of identifying a polymorphic variant which alters transcription of cytochrome CYP3A5, the method comprising providing a DNA construct having a sequence encoding a reporter molecule operably linked to a DNA fragment comprising said transcription regulatory region, and introducing said construct into a cell and monitoring for the level of expression of said reporter molecule.

33. A method of identifying transcription factors capable of hybridising to a DNA sequence from a transcription regulatory region adjacent to DNA encoding cytochrome CYP3A5, said method comprising contacting said DNA sequence including said transcription regulatory region with potential transcription factors and identifying any transcription factor complexed to said DNA sequence.

34. A method of identifying compounds acting on a transcription regulatory region of a DNA sequence

encoding CYP3A5, the method comprising transforming a cell with a DNA construct comprising the sequence of said regulatory region, and which regulatory region is operably linked to a sequence encoding a reporter molecule, contacting said cell with a test compound and identifying any altered expression of said reporter molecule.

35. A method of purifying transcription factors from a sample which are capable of binding to DNA from a transcription regulatory region of a sequence encoding cytochrome CYP3A5, the method comprising contacting a DNA sequence including said transcriptional regulatory region with a mixture of transcription factors and identifying any complexes of said transcription factors and said sequence.

36. A method according to any of claims 32 to 35 wherein said transcription regulatory region includes a mutation in a recognition site for a transcription factor of said regulatory region.

37. A method according to any of claims 32 to 36 wherein said mutation occurs in an activator protein-3 motif (AP-3) and/or a basic transcription element (BTE).

38. A method according to any of claims 36 or 37 wherein said mutation occurs at any one of positions -475 or -147 of the transcription regulatory region adjacent to the sequence encoding CYP3A5, the sequence of which region is illustrated in Figure 7.

39. A method according to any of claims 32 to 38 wherein the transcription regulatory region comprises

the mutations T<sub>-475</sub>G and A<sub>147</sub>G.

difference = difference in threshold cycle  
between samples, as described in the methods  
section wt = samples with the wild type  
sequence in the 5' flanking region as  
previously published (11) Het = samples  
heterozygous for the linked polymorphisms,  
A-147G and T-475G.

10 Fig. 2a: is a diagram of relative position of  
primers, and of the recognition site for the  
restriction enzyme *Tai* I, which is  
introduced into the PCR product utilising  
mismatches primer 3A5R1 when the wild-type  
"A" nucleotide is present at position -147,  
15 and is lost when the mutant "G" nucleotide  
is present.

20 Fig. 2b: is a diagrammatical representation of expected restriction fragments for each possible genotype for the A-147G variant, i.e. homozygous wild-type, heterozygous and homozygous mutant.

25 **Fig. 2c:** is an illustration of a 1.5% agarose gel of  
Tai I restriction digest of 3A5F2/3A5R1 PCR  
product for detection of the A<sub>-147</sub>G variant.  
Lane 1. 100 bp ladder. Lanes 2 & 7.  
Reference undigested PCR products. Lane 3.  
30 Sample homozygous for the wild-type "A"  
nucleotide at position -147. Lanes 10, 11,  
16. Samples heterozygous for the A<sub>-147</sub>G  
variant.

35 Fig. 3a: is a diagram of relative position of

- 14 -

5

primers, and of the recognition sites for the restriction enzyme *Alu* I. The forward recognition site is introduced into the PCR product utilising mismatched primer 3A5F2 when the wild-type "T" nucleotide is present at position -475, and is lost when the variant "G" nucleotide is present.

10

**Fig. 3b:** is a diagrammatical representation of expected restriction fragments for each possible genotype for the T<sub>-475</sub>G variant, i.e. homozygous wild-type, heterozygous and homozygous mutant.

15

**Fig. 3c:** is an illustration of a 12.5% polyacrylamide ExcelGel of *Alu* I restriction digest of the 3A5F2/3A5R1 PCR product for detection of the T<sub>-475</sub>G mutation. Lane 1.100 bp ladder. Lanes 2, 5, 6, 7 & 8. Samples homozygous for the wild-type "T" nucleotide at position -147. Lanes 3, 4, 9. Samples heterozygous for the T<sub>-475</sub>G mutation. Fragment X - additional digestion product resulting from re-amplification of original template by primers 3A51/3A52.

20

25

30

**Fig. 4a:** is a comparison of 1-OHM/4-OHM metabolic ratios between samples with the linked mutations (HET group) and those wild-type for the mutations at positions -147 and -475 (WT group). Mean and quartiles are shown for each group, as is overall mean for the combined groups (central line).

35

**Fig. 4b:** is a comparison of CYP3A5 expression (ln

5

transformed) between samples with the linked mutations (HET group) and those wild-type for the mutations at positions -147 and -475 (WT group). Mean and quartiles are shown for each group, as is overall mean for the combined groups (central line).

10

**Fig. 5:** is a Western blot analysis of CYP3A5 protein expression in liver samples. A Western blot of microsomes prepared from liver samples and probed with a CYP3A5 specific antibody. Liver samples containing the linked polymorphisms at -147 and -475 (wt group) are marked \* (sizes indicated in kDa from Wide Range Colour Marker (signs)).

15

20

**Fig. 6:** is a list of oligonucleotide mismatch primers used in accordance with the invention, where the underlined nucleotide indicates the sequence mismatch.

25

**Fig. 7:** is an illustration of the nucleotide sequence of the 5' flanking region relative to the DNA sequence encoding CYP3A5.

30

35

**Fig. 8:** is an illustration of the results obtained from an Electrophoretic mobility shift assay (EMSA) of A<sub>-147</sub>G oligonucleotide. EMSA was carried out as described in materials and methods. Lane 1 : A-147G oligonucleotide without HeLa nuclear extract; lanes 2-8 : in the presence of HeLa nuclear extract; lanes 3 and 4 : in the presence of 50 - 100 fold molar excess of unlabeled A-147G oligonucleotide; lanes 5 and 6; in the

- 16 -

presence of 50 - 100 fold molar excess of unlabeled wild type oligonucleotide; lanes 7 and 8 : in the presence of 1 and 2 microlitres of anti-Spl antibody.

5

**Fig. 9-9d:** are illustrations of the results obtained from the 'find patterns' program of the GCG sequence analysis package.

10

#### Experimental Procedures

##### *Liver microsome preparation*

15 Human liver samples were obtained from kidney transplant donors, and flash-frozen immediately on removal. Human liver microsomes were prepared according to previously described protocols (21), and protein content was determined by the method of Lowry 20 as modified by Miller (22).

##### *Midazolam hydroxylase assay*

25 The rates of midazolam overall metabolism and of the formation of 1- and 4-OH-midazolam were determined as follows. Each incubation vessel contained an aliquot of the microsomal suspension (containing 1 mg of 30 microsomal protein) in 1.15 % KCl - 0.01 M phosphate buffer pH 7.4; 10  $\mu$ l of a stock solution of 6 mM midazolam dissolved in DMSO to reach a final midazolam concentration of 60  $\mu$ M; 500  $\mu$ l of a co-factor mixture containing 0.5 mg of glucose-6-phosphate, 0.5 mg of MgCl<sub>2</sub>.6H<sub>2</sub>O, 0.5 units of glucose-6-phosphate dehydrogenase dissolved in 0.5 M Na-K-phosphate 35 buffer, pH 7.4 and a 1.15 % KCl - 0.01 M phosphate

buffer pH 7.4 to bring the incubation volume to 0.9 ml. After a pre-incubation for 5 min at 37°C, the incubations were started by adding 100  $\mu$ l of a solution of 1.25 mg/ml NADP to reach a final 5 concentration of 0.125 mg/ml. Tubes were continuously shaken at 100 oscillations/min in an Heto shaking waterbath. Blank incubates with boiled microsomes were incubated under identical conditions as the control incubates. The incubations were stopped after 10 30 min by immersing the tubes in dry ice. Samples were stored at  $\leq$  -18°C until analysis. The incubation samples were analysed for unchanged midazolam and for its metabolites 1'- and 4-hydroxymidazolam by HPLC with UV-detection.

15

***HPLC determination of midazolam metabolites***

The 1-ml samples of midazolam were thawed and diluted with 1 ml DMSO. Samples were sonicated for 10 min, 20 centrifuged and an aliquot of the supernatant was injected directly onto the HPLC-column. The HPLC apparatus consisted of a Waters 600 MS pump. The samples were injected automatically, using a WISP 717 plus automatic injector. Stainless steel columns (30 25 cm x 4.6 mm i.d.) were packed with Kromasill 18 (5  $\mu$ m) bound phase by a balanced density slurry procedure (Haskel DSTV 122-C pump, 10<sup>7</sup> Pa). UV-detection at 230 nm was performed using a Waters 996 Diode Array Detector. Elution at 1-ml/min started with a short 30 gradient from 100% 0.1 M ammonium acetate, pH 7.0 (solvent system A) to 50% of solvent system A and 50% of solvent system B containing 1M ammonium acetate pH 7.0, methanol and acetonitrile (10/45/45), over a 1-min period, followed by a second gradient to 100% 35 solvent system B in 15 min. This solvent composition

was held for 2 min before equilibration with the starting conditions. The identity of the metabolites of midazolam was confirmed using mass spectroscopy. The conversion of UV-peak areas into ng was performed 5 by a Millennium 2020 CDS system on a calibration curve of midazolam. This calibration curve was made up after injection of known amounts of the drug (0, 1059, 2117, 3176 and 5028 ng) and linear (weighted by 1/x) regression analysis of the corresponding UV-peak 10 areas. The equation of the calibration curve was  $ng = 0.000333 \times \text{area}$  ( $r^2 = 0.9997$ ,  $n = 5$ ). The metabolic activity was expressed as pmol metabolite formed/min mg protein, and a metabolic ratio was determined for each sample according to the ratio of 10HM/40HM in 15 each sample.

***Genomic DNA preparation***

20 DNA was isolated from frozen liver samples using a QIAamp Tissue Kit (QIAGEN) in accordance with the manufacturer's instructions.

***RNA preparation***

25 RNA was isolated from the liver samples using a QIAGEN RNAeasy Midi Kit (QIAGEN), according to manufacturers instructions. Twenty  $\mu$ g of RNA was treated with RNase-free DNase I (Boehringer Mannheim), for 30 min at 37°C in 20 mM Tris-HCl, pH 8.0, 100 mM MgCl<sub>2</sub>. Samples were 30 phenol/chloroform extracted, precipitated and resuspended in 30  $\mu$ l of TE buffer. Two and a half  $\mu$ g of the treated sample was reverse transcribed for 50 minutes at 42°C in 1 x first strand buffer, 0.01M DTT and 0.5M dNTPs using 0.5  $\mu$ g of oligo(dt) random 35 primers and 200 units SuperScript II Reverse

Transcriptase (GibcoBRL) for use on the ABI Prism 7700 Sequence Detection System (SDS).

*Sequencing of the CYP3A5 5' flanking region*

5

A 1343 bp 5' flanking region of CYP3A5 was PCR amplified from genomic DNA isolated from liver samples, using primers 3A51 (5'-

GGAAAGCAACCTACATGTCCATC) and 3A52 (5'-

10 ATCGCCACTTGCCTTCTTC) based on the published sequence of Jounaidi *et al.* (11). PCR conditions were 1 cycle of 95°C for 1 min, 30 cycles of 95°C for 1 min, 57°C for 30 sec, 72°C for 2.5 min, and 1 cycle of 72°C for 10 min. PCR products were purified using a QIAquick 15 PCR Purification Kit (QIAGEN), sequencing primers were designed (Table 1), and used to directly sequence the PCR product on both sense and antisense strands by cycle sequencing using the ABI BigDye Terminator cycle sequencing kit (Perkin Elmer). Sequencing reactions 20 were analysed on an ABI 377 automated sequencer.

Contig sequences were aligned and compared using the Sequence Editor version 1.0.3 software packages (Perkin Elmer) and manually edited for identification of heterozygote positions.

25

*PCR detection assays for the A-147G and T-475G mutations*

All PCR assays were performed utilising a 1 in 100 dilution of the original 3A51/3A52 PCR product as 30 template, under the following conditions: 1 cycle of 95°C for 1 min, 30 cycles of 95°C for 1 min, 55°C for 30 sec, 72°C for 1 min, and 1 final cycle of 72°C for 10 min. All products were sequenced to confirm the identity of the product as CYP3A5. Oligonucleotide 35 mismatched primers utilised in the assays were: 3A5F1

- 20 -

(5'-GGGTCTGCTGGCTGCGC), 3A5F2 (5'-  
GGGGTCTGTCTGGCTGAGC), and 3A5R1 (5'-  
TTATGTGCTGGAGAAGGACG), where positions of mismatches  
are underlined.

5

For the A<sub>475</sub>G mutation, PCR was performed using primer  
pair 3A5F2 and 3A5R1. Twenty  $\mu$ l of PCR product was  
digested for a minimum of 3 hours at 65°C using 15  
units of *Tai* I, and the restriction fragments  
10 visualised by ethidium bromide staining after  
electrophoresis on a 1.5% agarose gel.

For the T<sub>475</sub>G mutation, PCR was performed using primer  
pair 3A5F2 and 3A5R1 as described above. Twenty  $\mu$ l of  
15 PCR product was digested with 15 units of *Alu* I for a  
minimum of 3 hours, and restriction fragments were  
separated by electrophoresis on a 12.5% ExcelGel on a  
Pharmacia Multiphor Electrophoresis system  
(Pharmacia). Fragments were visualised by silver  
20 staining in a Hoeffer Automatic Gel Stainer  
(Pharmacia).

To detect the presence of mutations on the same  
chromosome, PCR was performed using primers 3A5F1 and  
25 3A5R1. Twenty  $\mu$ l of PCR product was digested for a  
minimum of 3 hours at 65°C using 15 units of *Mvn* I,  
and the resulting restriction fragments were  
visualised by ethidium bromide staining after  
electrophoresis on a 1.5% agarose gel.

30

***Relative quantification and comparison of CYP3A5 RNA***

Relative levels of CYP3A5 mRNA were determined by real  
time PCR using the ABI 7700 SDS (Perkin Elmer).  
35 Optimal primers and probes for the detection of CYP3A5

- 21 -

were designed using the PrimerExpress program, and subsequently checked to ensure specificity for CYP3A5. Primers utilised for the quantification PCR were:  
5 forward - 5'-AAGTGGCGATGGACCTCATC-3'; reverse - 5'-GAGGAGCACCAGGCTGACA-3'. The TaqMan probe was labelled with the 5' reporter dye 6-carboxy-flouresine (FAM), and had the sequence 5'-CAAATTTGGCGGTGGAAACCTGGC-3'. Optimal primer/probe ratios and concentrations were determined and the experiments run according to 10 standard protocols for the ABI 7700 Standard Detection System. CYP3A5 mRNA expression for all samples was normalised against the expression of  $\beta$ -actin mRNA. The threshold cycle (Ct) is the PCR cycle number where the ABI 7700 begins to detect an increase in fluorescent 15 signal associated with the linear amplification of PCR product. The Ct value is dependent on the initial amount of template copy. Quantities of CYP3A5 in each sample were determined by averaging the Ct from 3 separate PCR reactions of each sample. Relative 20 differences in Ct between samples were calculated by subtracting the Ct of each sample from the highest Ct within the samples (lowest expression). Since the amount of PCR product doubles with every cycle in the linear range of a PCR the differences in Ct were 25 converted into estimated differences of mRNA quantity between the samples by calculating  $2^{\delta\text{Ct}}$ , where  $\delta\text{Ct}$  is the difference in cycle threshold between two samples.

Negative controls were performed on each run to ensure 30 that no signals were due to DNA contamination. Control samples consisted of RNA samples which had been treated in exactly the same manner as for the quantitative PCR, but without the addition of the reverse transcriptase.

***Statistical Analysis***

Statistical analysis was performed on the JMP Statistical program version 3.2.2 (SAS Institute Inc.). Metabolic ratio and CYP3A5 mRNA expression data were checked to ensure that they conformed to a normal distribution. CYP3A5 mRNA expression data did not conform to a normal distribution and were ln-transformed, afterwhich the data was normally distributed. Metabolic ratios and expression levels were compared between groups using a t-Test.

***Western Blot Analysis***

Forty micrograms of microsomal protein prepared from each liver were solubilised in an equal volume of Laemmli sample buffer (Biorad) by four cycles of freezing and boiling for 10 minutes. Samples were loaded onto pre-cast 10% SDS-PAGE Ready Gels (Biorad) and electrophoresed for 1 hour at 180 V. Separated proteins were transferred onto Hybond-P membranes (Amersham) using a Trans-blot SD apparatus (Biorad). Membranes were blocked by an overnight incubation at 4°C in 1x PBS containing 5% (w:v) nonfat milk and 0.1% (v:v) Tween. Membranes were incubated at ambient temperature for 1 hour in a 1:3000 dilution of specific human CYP3A5 antibody (Gentest) in 1X PBS, 2.5% nonfat milk, then rinsed four times in 1x PBS, 2.5% (w:v) nonfat milk, 0.1% (v:v) Tween. Membranes were incubated at ambient temperature for 1 hour in a 1:5000 dilution of Anti-Rabbit IgG peroxidase conjugate (Sigma) in 1x PBS, 2.5% (w:v) nonfat milk, and rinsed as previously. The membranes were developed using the ECL Plus Western Blotting Detection System (Amersham) according to

manufacturer's instructions, and visualised by autoradiography using Kodak X-Omat film (sigma).

Example 1

5

*Midazolam phenotyping*

A panel of 39 liver samples was phenotyped for CYP3A5 activity, using the metabolism of midazolam to its 1-OH metabolite as a marker of activity. Human liver 10 microsomal samples containing CYP3A5 in addition to CYP3A4 exhibit a significantly greater ratio of 1-OHM to 4-OHM compared with samples containing only CYP3A4. 15 1-OHM/4-OHM ratios between 5 and 9 were observed for microsomes containing both CYP3A4 and CYP3A5. Samples containing only CYP3A4 showed 1-OHM/4-OHM ratios < 4 (15). Analysis of the CYP3A5 phenotypes in our data set showed a clear bimodal distribution, with 6 samples (15%) having metabolic ratios greater than 5, 20 and the remaining samples having metabolic ratios lying between 1.5 and 3.5 (see Fig. 1a). Of the 39 liver samples from which microsomes were prepared for metabolic analysis, sufficient tissue was available 25 for full DNA and RNA analysis for 26, which included 6 samples lying in the higher metabolic ratio range. In addition to these 26 samples microsomes for protein analysis were available for a further 3 samples, all of which had metabolic ratios of <4.

30 *Analysis of CYP3A5 gene 5' flanking region*

The 5' flanking region of CYP3A5 was PCR-amplified from genomic DNA of all 26 samples and sequenced in full, as shown in Figure 7. Alignment showed that the 35 region was well conserved. Only a small number of

inter-individual variations were identified in addition to a few variations from the published sequence (Table 2.). All variants detected were heterozygous, and all samples heterozygous for the 5 more frequent A<sub>-147</sub>G mutation were also heterozygous for the T<sub>-475</sub>G mutation, suggesting that the two mutations were linked. These two mutations fall within two separate putative regulatory elements, a basic transcription element (BTE: A<sub>-147</sub>G) and an 10 activator protein-3 motif (AP-3: T<sub>-475</sub>G). None of the remaining variants fell within putative regulatory domains.

PCR assays were developed to confirm the presence of 15 the A<sub>-147</sub>G and T<sub>-475</sub>G mutations individually, and to ascertain if the two mutations were on the same, or on separate chromosomes. The PCR assay for the A<sub>-147</sub>G mutation was based on the creation of a recognition site for the restriction enzyme *Tai* I by utilising an 20 oligonucleotide mismatch primer (3A5R1). This primer introduces a *Tai* I recognition site only when the wild-type "A" nucleotide is present at position -147. Digestion of the 369bp product with *Tai* I yields fragments of 349 and 20bp for the wild-type sequence, 25 whilst the product remains undigested if the mutant "G" nucleotide is present (Fig. 2). Similarly, for the detection of the T<sub>-475</sub>G mutation a second oligonucleotide mismatch primer was used (3A5F2). This primer introduces a recognition site for the 30 restriction enzyme *Alu* I when the wild-type T nucleotide is present at position -475, digesting the product to yield fragments of 318, 33 and 18 bp. This site is lost when the mutant G nucleotide is present, yielding digestion products of 336 and 33 bp (Fig. 3).

To determine if the mutations were present on the same chromosome a PCR assay was developed utilising two oligonucleotide mismatch primers (3A5F1 and 3A5R1), both primers introducing recognition sites for the 5 restriction enzyme *Mvn* I when the mutant nucleotides are present at positions -147 and -475. If the mutations are present on the different chromosomes then the original 369 bp product is digested to yield products of 349/350 bp and 20/19 bp (inseperable by 10 gel electrophoresis), whilst if present on the same chromosome the fragment is digested to yield products of 330 and 20/19 bp (data not shown). In addition to confirming the individual genotypes of the samples as determined by sequencing the two mutations were, in 15 all cases, linked on the chromosome (data not shown).

***Relationship between CYP3A5 allelic variants, CYP3A5 mediated metabolism, CYP3A5 mRNA and protein expression***

20 Samples were grouped according to genotype: "Wild-type" or "mutant" (containing the linked polymorphisms), and the 1-OHM/4-OHM metabolic ratios (mr) were compared between the groups (Fig. 4a). With 25 the exception of one outlier (liver sample number, mr = 2.08), all individuals carrying the linked mutations had metabolic ratios > 5.0, whilst the wild type group all possessed metabolic ratios of < 3.5. The mean metabolic ratios for the mutant group were 30 significantly higher than those from the wild-type group ( $6.0 \pm 2.0$  versus  $2.7 \pm 0.42$ , mean  $\pm$  standard deviation;  $p < 0.001$ ).

35 Quantitative PCR was used to ascertain if the mutations in the 5' flanking region were affecting

gene expression. Whilst mRNA levels showed greater variation than the metabolic data, a degree of bimodality was observed (Fig. 1b). The mutant group had CYP3A5 mRNA levels skewed towards the higher end 5 of the expression range, showing significantly higher levels of CYP3A5 mRNA than the wild type group (mean  $\ln Ct = 4.03$ , standard deviation = 0.97, against mean  $\ln Ct = 2.06$ , standard deviation = 1.2,  $p < 0.006$ ) ((Fig. 4b). In this case the outlier (presenting with 10 the mutant genotype, but wild type metabolic ratio) also fell within the lower range of expression ( $\ln Ct = 2.9$ ).

15 The level of CYP3A5 protein expression levels was determined for 29 liver samples by Western blot analysis using a CYP3A5 specific antibody. A single band of 52 kDa corresponding to CYP3A5 was clearly apparent in some samples. With the exception of the single outlier with the high expression genotype 20 (mutant) and low metabolic ratio phenotype (wild-type), all samples which possessed the high expression genotype, a high metabolic ratio and high RNA expression level clearly show high levels of CYP3A5 expression when compared to those samples with the low 25 expression genotype and phenotype (Fig. 5). The single outlier with the high expression genotype, but low expression phenotype showed levels of CYP3A5 expression similar to those in the low expression genotype group. Longer exposure of the Western blot indicated that a very low level of CYP3A5 expression 30 was apparent in most samples (data not shown).

35 The 5' flanking sequences of CYP3A5 obtained in this study are virtually identical to those published by Jounaidi *et al.* (11), and show little inter-individual

variation in sequence. Interestingly, Jounaidi *et al.* sequenced two human genomic clones, one of which contained the two linked mutations described in detail in this report. This would suggest that one clone was 5 derived from an individual in the low expression group, and one from an individual in the high expression/metabolism group.

Previous studies had suggested that CYP3A5 was 10 expressed in 10-30% of livers (7, 8, 9) whilst another study has stated that some expression is constitutive in all samples (10). The present study supports the findings that some CYP3A5 expression is constitutive, with some metabolic activity and mRNA being detected 15 in all livers studied, although CYP3A5 protein was not convincingly demonstrated in all samples using the procedures required. We detected enhanced RNA and protein expression in 23% of the samples for which tissue was available (6 out of 26), which is similar 20 to the fraction of liver showing expression in previous studies. This supports the finding of Boobis *et al.* (10) that some show low level expression is constitutive in all liver samples although this can only be detected using more sensitive detection 25 techniques (such as PCR, and not by Western or Northern blot analysis).

Whilst both polymorphisms detected lie within putative transcriptional regulatory elements, we suspect that 30 the variant within the BTE is more likely to be responsible for altered expression since it has been reported that a BTE flanking the TATA box accounts for the constitutive expression of CYP1A1, and a similar region has been found in several other CYP genes 35 including CYP2B1, CYP2B2, CYP2E1 (16) CYP3A4 (13) and

CYP3A7 (12). In the case of CYP3A4 gene this element has been shown to bind nuclear extracts (13) and a basic transcription element binding factor for CYP3A7 (12), pointing to a role of this region in the general control of cytochrome P450 expression. The exact mechanism of up-regulation of CYP3A5 expression in the allelic variant described here remains to be determined although the presence of one of the mutations within the BTE, and the relevance of this element for the expression of other P450s indicates a possible mechanistic link. Using methylation interference footprinting, it has been shown that all guanine residues within the BTE, and other guanine residues in the vicinity, interacted with the transcriptional factor Sp1 (19). Given that the mutation within the BTE (Sp1) described herein alters an adenine residue to a guanine residue, then this could facilitate binding of transcription factors to the variant form of the Sp1.

Although there is considerable overlap in the range of CYP3A5 mRNA levels seen in the homozygous and heterozygous group, the distribution of metabolic ratios is clearly bimodal, as is the amount of CYP3A5. We cannot exclude the presence of other polymorphisms that may affect the translation efficiency or protein stability of CYP3A5. But given the better correlation between DNA polymorphism and protein level and the notorious liability of RNA, the simpler explanation is that differential RNA degradation or yield (due to differences in sample handling) has blurred the distinction between high and low expressers. Whatever the explanation for the discrepancy at the mRNA level, it does not in any way diminish the predictive value of the DNA polymorphism described.

There is, however, one individual whose genotype (heterozygous mutant) is not predictive of his metabolic phenotype (low expression). The fact that CYP3A5 protein as well as mRNA levels were low in this 5 outlier indicates that the explanation must be sought at the transcriptional level, e.g. in transcription factors controlling CYP3A5 expression.

An AUG element in the 5'- untranslated region of the 10 BTEB gene has been shown to be, at least in part, responsible for cell specific translational control of BTEB (20). Mutations within this region were shown to affect BTEB translation. Therefore, whilst the outlier in our study has a high expression genotype for CYP3A5 15 expression, this individual may have a "poor" expression phenotype for BTEB. Additionally, it is possible that a mechanism similar to that responsible for inducing CYP1A1 expression may also affect CYP3A5 expression. In addition to the BTE, CYP1A1 expression 20 is mediated by a xenobiotic responsive element (XRE). In this case inducers enhance expression by binding to a cytosolic receptor (Ah receptor) which is translocated into the nucleus (possibly in association with an accessory protein coded for at the Arnt gene), 25 where it binds the XRE (17, 18). Although variations in these and other transcription factors could further modulate CYP3A5 expression, this does not detract from the fact that the polymorphism described here seems to be the major determinant of CYP3A5 expression, at 30 least in liver.

Despite the relatively small number of samples 35 available for analysis in the present study, strong associations have been found between the two linked polymorphisms on the one hand and both expression and

- 30 -

CYP3A5 mRNA, protein and activity levels in the liver on the other hand. The unravelling of a genetic mechanism for the polymorphic metabolism by CYP3A5 will have important consequences in the field of pharmacogenetics. The ability to predict metabolism by genotyping will greatly facilitate disease association studies and may also help to explain adverse reactions or poor response to therapeutics which are metabolised by this cytochrome P450 isoform. It will also help in delineating which factors affecting CYP3A5 activity are genetic and which are environmental; for both further work will be required to fully understand the complex variation in expression observed with this enzyme.

15

#### **Putative promoter sequence analysis**

##### **Materials and Methods:**

The sequence of the regulatory region of CYP3A5 was analysed with the 'findpatterns' program of the GCG sequence analysis package (GCG, Madison, Wisconsin). This program finds specific DNA sequence motifs, patterns, and transcription binding sites, whose sequences are stored in the program, and are present in the sequence of interest. In the present analysis, at most one single mismatch or error per pattern is allowed in the sequence of interest, to detect if the two reported variations alter any known motifs or transcription binding sites. Results are identified in Figures 9 to 9d.

The first, GCGTG to GCTTG variation removes binding sites for MBF-I\_CS, MRE\_CS2, and CNBP-SRE.

35

- 31 -

The second, CC~~A~~CC to CCGCC variation  
replaces binding sites for apoE-undefined-site-3,  
ApoE\_B1, APRT-CHO\_US, and APRT-human\_US  
by GCF-consensus, APRT-mouse\_US, GC-box\_(1),  
5 DSE\_(1), Sp1\_CS4, Sp1-hsp70\_(1), hsp70.2, Sp1-IE-  
3.3, Sp1-IE-4/5, IRE\_(1), Sp1-TPI\_(4)  
does not affect the Yi-consensus pattern

Both mutations affect transcription factor binding  
10 sites.

#### **Electrophoretic mobility shift assay (EMSA)**

An EMSA was carried out using the Sp1 NUSHIFT Kit from  
15 Geneka Biotechnology Inc. (Montreal, Canada) according  
to the manufactures instructions. Briefly, a 31-mer  
double-stranded oligonucleotide corresponding to the  
CYP3A5 5'-untranslated region containing the A-<sub>14</sub>G  
polymorphism (5'-GGC AGC TGC AGC CCC GCC TCC TTC TCC  
20 AGC A-3') was end-labeled with <sup>32</sup>-P using T4  
polynucleotide kinase. 50,000 cpm (0.5 ng)  
oligonucleotide was incubated with 2 µg HeLa nuclear  
extract for 30 min at 16°C. Unlabeled mutant or  
wildtype (5-'GGC AGC TGC AGC CCC ACC TCC TTC TCC AGC  
25 A-3') oligo nucleotide was added in 50-fold or 100-  
fold excess as indicated. 1 or 2 µl anti-Sp1 rabbit  
polyclonal antibody was pre-incubated with the nuclear  
extract at 4°C for 30 min as indicated. Nuclear  
extract, anti-Sp1 antibody and binding buffers were  
30 from Geneka Biotechnology Inc. Samples were separated  
on a 5% polyacrylamide (39:1) gel, in TGE buffer (25  
mM Tris, 190 mM glycine, 1 mM EDTA, pH 8.3). The dried  
gel was exposed to X-ray film.

**RESULTS**

Analysis of the 5'-untranslated region of the CYP3A5 gene indicated that the A<sub>-147</sub>G polymorphism might create a binding site for the transcription factor 5 Sp1. An electrophoretic mobility shift assay (EMSA) was carried out to test this hypothesis. An oligo nucleotide containing the A<sub>-147</sub>G polymorphism was used to assay for binding factors present in HeLa nuclear extracts. A band shift was observed (Figure 8, lane 2) 10 which was competed away with 50- and 100-fold excess respectively of unlabeled oligo nucleotide (Figure 8, lanes 3 and 4), but not with wildtype oligo nucleotide (Figure 8, lanes 5 and 6). This clearly indicates the presence of a protein factor in HeLa nuclear extracts 15 capable of binding to the A<sub>-147</sub>G polymorphism region, but not to the wildtype region. Incubations in the presence of an antibody specific for the transcription factor Sp1 resulted in supershifting of the A<sub>-147</sub>G polymorphism oligo nucleotide (Figure 8, lanes 7 and 20 8), indicating that Sp1 is binding to the A<sub>-147</sub>G polymorphism site.

This change in binding affinity of transcription factor Sp1 to the 5'-untranslated region of the CYP3A5 gene might account for the increase in transcription 25 from the A<sub>-147</sub>G polymorphic promoter and in turn, might contribute to the increase in metabolic rates correlated with the A<sub>-147</sub>G polymorphisms.

**Genotyping of the cytochrome expression**

30 A group of 300 healthy Caucasian volunteers was genotyped for variations T<sub>-475</sub>>G and A<sub>-147</sub>>G of the cytochrome P450 3A5 gene.

Test rationale

The first objective concerned allele/genotype frequencies.

5 Because the initial study included only 30 to 35 different individuals, allele/genotype frequencies could not be determined. Genotyping a group of 300 subjects should permit determination of these frequencies and to check whether they are in agreement  
10 with the Hardy-Weinberg equilibrium.

10 The second objective concerned the linkage of the two variations. In the initial study, all samples with the gene variations T-<sub>475</sub> > G and A-<sub>147</sub> > G (only 6 in total) were linked. To verify the suggested linkage, both of  
15 these CYP 3A5 polymorphisms were genotyped on a larger population.

Materials and methods

20 In order to minimize genotyping errors, genomic DNA samples from 300 healthy Caucasian volunteers were genotyped in a microtiterplate based format, which ensured a blind and completely independent duplicate analysis of each individual sample.

25 A 1343 bp 5' flanking region of CYP3A5 was PCR-amplified from genomic DNA using primers 3A51/3A52. PCR assays for both variations were performed using a 1/100 dilution of the original 3A51/3A52 PCR product as template. Mismatch primers 3A5F2 and 3A5R1 were utilised for both assays. For the A-<sub>147</sub> > G mutation the PCR product was digested with restriction enzyme *Tai* I, and for the T-<sub>475</sub> > G mutation the PCR product was digested with restriction enzyme *Alu* I. After digestion the restriction fragments were separated by polyacrylamide gel electrophoresis and visualised by  
30  
35

silver staining. The genotypes were determined based on the DNA fragment patterns by two independent observers.

5      **Results**

1. Allele/genotype frequencies

In the population of 300 individuals, 53 heterozygous subjects (18%) were carrying one copy of each of the variations, 246 subjects (82%) were homogenous for 10 A-<sub>147</sub> and T-<sub>475</sub>, and one individual (0.3%) was carrying variations G-<sub>147</sub> and G-<sub>475</sub> on both alleles (homozygous). These frequencies are in agreement with 3A5 expression found in previous studies (7,8,9). The allele frequencies are in agreement with the 15 Hardy-Weinberg equilibrium (Table 3).

2. Linkage of variations T-<sub>475</sub> > G and A-<sub>147</sub> > G

In all individuals, respectively variations T-<sub>475</sub> and A-<sub>147</sub>, and variations G-<sub>475</sub> and G-<sub>147</sub>, were equally represented in genotypes, indicating a strong linkage 20 between both variations. Whether this linkage between both variations has some functional significance needs to be clarified further. As a consequence of the linkage, future genotyping will require only the analysis of one of the variations, whether it is the 25 functional variant or not.

References

1. Cholerton, S., Daley, A.K., Idle, J.R., The role of individual human cytochrome P450 in drug metabolism and clinical response. *Trend Pharmac. Sci.* **13**, 434-439 (1992).
2. Shimada, T., Guengerich, F.P. Evidence for cytochrome P-450NF, the nifedipine oxidase, being the principal enzyme involved in the bioactivation of aflatoxins in human liver. *J. Biol. Chem.* **86**, 462-465 (1989).
3. Kormori M. et al. Fetus specific expression of a form of cytochrome P-450 in human livers. *Biochemistry* **29**, 4430-4433 (1990).
4. Hoyo-Vadillo, C. et al. Pharmacokinetics of nifedipine slow release tablets in Mexican patients: further evidence for an oxidative polymorphism. *J. Clin. Pharmac.* **29**, 816-820 (1989).
5. Renwick, A.G., Robertson, D.R.C., Macklin, B., Challenor, V., Waller, D.G., George, C.F. The pharmacokinetics of oral nifedipine - a population study. *Br. J. Clin. Pharmacol.* **25**, 701-708 (1988).
6. Schellens, J.H.M., Soons, P.A., Breimer, D.D. Lack of nifedipine plasma kinetics in a large population of healthy subjects. *Biochem. Pharmacol.* **37**, 2507-2510 (1988).
7. Aoyama, T., et al., Cytochrome P450 hPCN3 a novel cytochrome P450IIIA gene product that is differentially expressed in adult human liver. *J. Biol.*

Chem. 264, 10388-10395 (1989).

8. Wrighton, S.A. et al. Studies in the expression and metabolic capabilities of human liver cytochrome P450IIIA5 (HLp3). Mol. Pharmacol. 38, 207-213 (1990).

5

9. Schuetz, J., Beach, P., Guzelian, P.S. Selective expression of cytochrome P450 CYP3A mRNAs in embryonic and adult human liver. Pharmacogenetics, 4, 11-20 (1994).

10

10. Boobis A.R., Edwards, R.J., Adams, D.A., Davies, D.S. Dissecting the function of cytochrome P450. Br. J. Clin. Pharmacol. 42: 81-89 (1996).

15

11. Jounaidi, Y., Guzelina, P.S., Maurel, P., Vilarem, M.J. Sequence of the 5'-flanking region of CYP3A5 comparative analysis with CYP3A4 and CYP3A7. Biochem. Biophys. Res. Commun. 3, 1741-1747 (1994).

20

12. Itoh, S., et al. Genomic organisation of human fetal specific P-450IIA7 (cytochrome P-450HFLa)-related gene(s) and interaction of transcriptional regulatory factor with its DNA element in the 5' flanking region. Biochimica et Biophysica Acta. 1130, 133-138 (1992).

25

13. Hashimoto, H., et al. Gene structure of CYP3A4, an adult specific form of cytochrome P450 in human livers and its transcriptional regulation. Eur. J. Biochem. 218, 585-595 (1993).

30

14. Barwick JL., Quattrochi, L.C., Mills, A.S., Potenza, C., Tukey, Guzelian, P.S. Trans-species gene transfer for analysis of glucocorticoid-inducible

35

transcriptional activation of transiently expressed human CYP3A4 and rabbit CY{3A6 in primary cultures of adult rat and rabbit hepatocytes. *Molecular Pharmacology*, **50**, 10-16 (1996).

5

15. Gorski, J.C., Hall, S.D., Jones, D.R., VandenBranden, M., Wrighton, S.A. Regioselective biotransformation of midazolam by members of the human cytochrome P450 (CYP3A) subfamily. *Biochem. Pharmacol.* **9**, 1643-1653 (1994).

10 16. Fujii-Kuriyama, Y., Imataka, H., Sogawa, K., Yasumoto, K.I., Kikuchi, Y. Regulation of CYP1A1 expression. *FASEB J.* **6**, 706-710 (1992).

15

17. Nebert, D.W. The Ah locus: genetic differences in toxicity, Cancer, mutation and birth defects. *Critical Reviews in Toxicology*, **20**, 153-174 (1989).

20

18. Hoffman, E.C., et al. Cloning of a factor required for activity of the Ah (dioxin) receptor. *Science*, **252**, 954-958 (1991).

25

19. Sogawa, K., Kikuchi, Y., Imataka, H., Fujii-Kuriyama Y. Comparison of DNA-binding properties between BTEB and Sp 1. *J. Biochem.* **114**, 605-9 (1993).

30

20. Imataka, H., Nakayama, K., Yasumoto, K., Mizuno, A., Fujii-Kuriyama, Y., Hayami M. Cell specific translational control of transcription factor BTEB expression. The role of an upstream AUG in the 5' untranslated region. *J. Biol. Chem.* **269**, 20668-73 (1994).

35

21. Lavrijsen K., Van Houdt, J., Thijs, D.,

- 38 -

Meuldermans, W., Heykants, J. Induction potential of  
antifungals containing an imidazole or triazole  
moiety. Miconazole and ketoconazole, but not  
itraconazole are able to induce hepatic drug  
metabolising enzymes of male rats at high doses.  
5 *Biochem. Pharmacol.* **35**, 1867-78 (1986).

22. Miller, G.L. Protein determination for large  
numbers of samples. *Anal. Chem.* **31**, 964 (1959).

10 23. Guengerich, F.P. Characterisation of human  
cytochrome P450 enzymes. *FASEB* **6**, 745-748 (1992).

15 24. Watkins, P.B. Drug metabolism by cytochromes P450  
in the liver and small bowel. *Gastroenterology*  
*Clinics of North America*, **21**, 511-526 (1992).

20 25. Haehner, B.D. et al. Bimodal distribution of  
renal cytochrome P450 3A activity in humans. *Mol.*  
*Pharmacol.*, **50**, 52-59 (1996).

25

30

35

- 39 -

Table 1. Primers used for sequencing 5' flanking region of CYP3A5 from PCR product 3A51/3A52 (see text).

Primer	Orientation#	Position*	Sequence (5'-3')
3A51	F	-1237- -1217	GGAAGCAACCTACATGTCCATC
3A5p01	F	-978- -963	AGTACAGGGAGCACAG
3A5p08	R	-917- -932	CACCTATTCAATTCCCTG
3A5p02	F	-698- -684	TGCTATCACCAACAGAC
3A5p07	R	-689- 704	GGTGATAGCAATAGAC
3A5p03	F	-364- -349	AGGATGTGTAGGAGTC
3A5p06	R	-417- -434	CCTCACACAGATGTAACC
3A5p04	F	-176- -161	TAAGAACTCAGGTTCC
3A5p05	R	-178- -194	CAGAAACTGAAGTGGAG
3A52	R	+105- +87	ATCGCCACTTGCCTTCTTC

# F = 5' to 3', R = 3' to 5'

\* Primer locations are based on CYP3A5 sequence data of Jounaidi et al (11)

20

25

30

- 40 -

Table 2.

Position	Variant Sequence	Percentage
-475	T-K (T or G) heterozygote	30.6% (11/36)
-147	A-R (A or G) heterozygote	30.6% (11/36)

10

TABLE 3.

15 Hardy Weinberg Equilibrium test

Test: CYP3A5 -45 A&gt;G

Population: CON-JRF-1

		Observed values		Expected values	
		N	freq	N	freq
20	genotype AA	246	0.820	247.5	0.825
	genotype AG	53	0.177	50.0	0.167
	genotype GG	1	0.003	2.5	0.008
	total	300	1	300	1

25 1.112 = Chi-square (Pearson)

0.292 = p-value

1 = d.f.

	N	freq	
30	Allele A	545	0.908
	Allele G	55	0.092

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 99/04380

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 7 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>SCHUETZ E G ET AL: "Modulators and substrates of P-glycoprotein and cytochrome P4503A coordinately up-regulate these proteins in human colon carcinoma cells."</p> <p>MOLECULAR PHARMACOLOGY, (1996 FEB) 49 (2) 311-8. , XP000906895 the whole document</p> <p>—</p> <p>—/—</p>	1-4, 7-10, 17-28, 32-37

 Further documents are listed in the continuation of box C. Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*&\* document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

4 May 2000

19/05/2000

## Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.  
Fax: (+31-70) 340-3016

Authorized officer

Molina Galan, E

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/GB 99/04380

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE MEDLINE 'Online! US NATIONAL LIBRARY OF MEDICINE (NLM), BETHESDA, MD, US</p> <p>YOKOI T ET AL: "Genetic polymorphism of drug metabolizing enzymes: new mutations in CYP2D6 and CYP2A6 genes in Japanese." retrieved from STN Database accession no. 1998428039 XP002136986 abstract &amp; NIPPON YAKURIGAKU ZASSHI. FOLIA PHARMACOLOGICA JAPONICA, (1998 JUL) 112 (1) 5-14. REF: 45 ,</p>	1-4, 7-10, 25-28, 32-37
X	<p>JOUNAIDI, YOUSSEF ET AL: "Sequence of the 5'-flanking region of CYP3A5: comparative analysis with CYP3A4 and CYP3A7" BIOCHEM. BIOPHYS. RES. COMMUN. (1994), 205(3), 1741-7 , XP002136985 cited in the application the whole document</p>	17-24
A	EP 0 463 395 A (HOFFMANN LA ROCHE) 2 January 1992 (1992-01-02)	
A	WO 95 30772 A (KAMATAKI TETSUYA ; OTSUKA PHARMA CO LTD (JP)) 16 November 1995 (1995-11-16)	
A	<p>YOKOI T ET AL: "Genetic polymorphism of drug metabolizing enzymes: new mutations in CYP2D6 and CYP2A6 genes in Japanese." PHARMACEUTICAL RESEARCH, (1998 APR) 15 (4) 517-24. REF: 63 , XP000906892</p>	

## INTERNATIONAL SEARCH REPORT

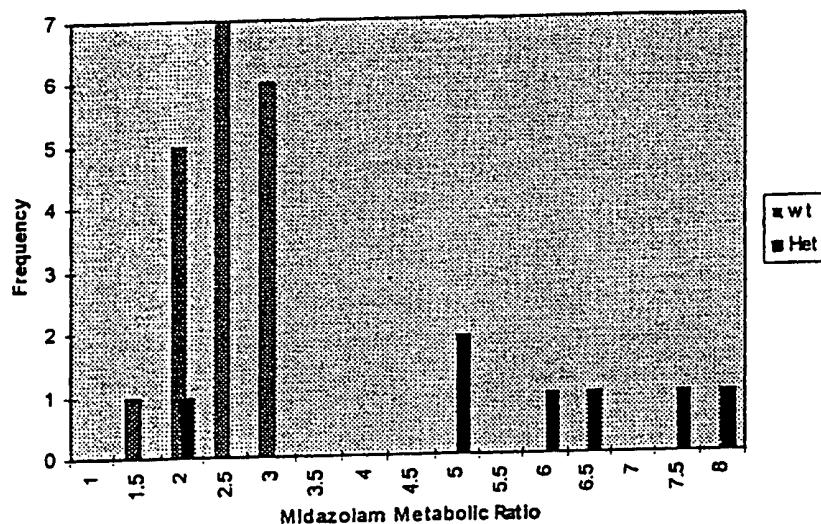
## Information on patent family members

International Application No

PCT/GB 99/04380

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP 0463395	A 02-01-1992	AU 645091 B		06-01-1994
		CA 2045012 A		23-12-1991
		JP 5211895 A		24-08-1993
		NZ 238610 A		25-11-1993
		US 5648482 A		15-07-1997
		US 5844108 A		01-12-1998
		DE 69126064 D		19-06-1997
		DE 69126064 T		28-08-1997
		ES 2103756 T		01-10-1997
WO 9530772	A 16-11-1995	JP 7298900 A		14-11-1995
		CA 2189638 A		16-11-1995
		CN 1151763 A		11-06-1997
		EP 0759476 A		26-02-1997

*FIG. 1a.*  
CYP3A5 Genotype/Phenotype Relationship



*FIG. 1b.*  
CYP3A5 mRNA Expression Related to Genotype

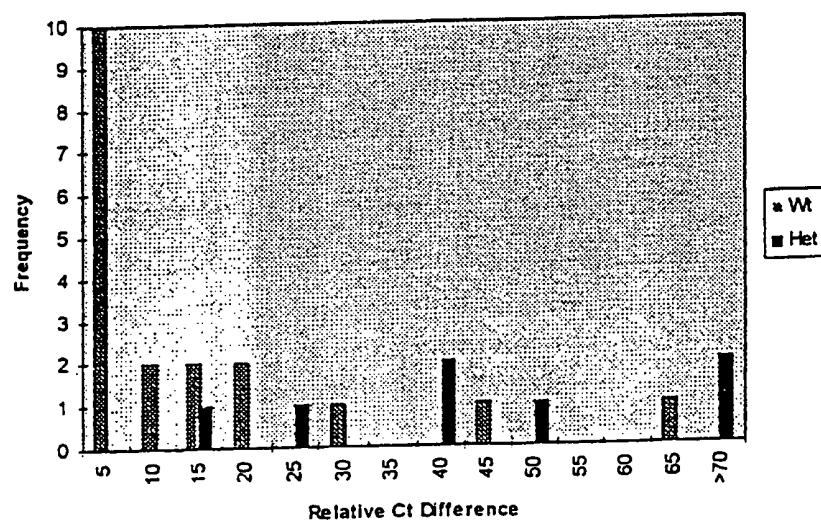
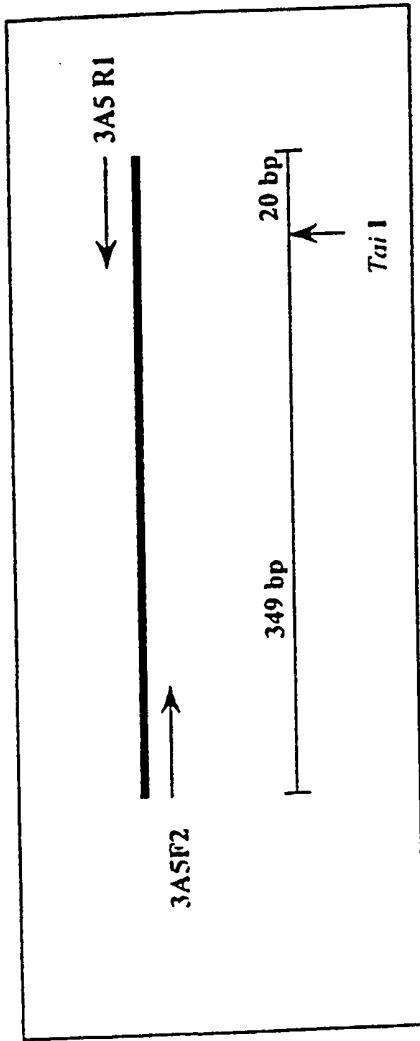
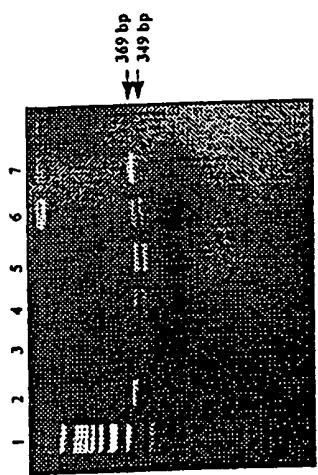


FIG. 2.

2a.



2c.



2b. Wt/Wt Wt-147G -147G/-147G

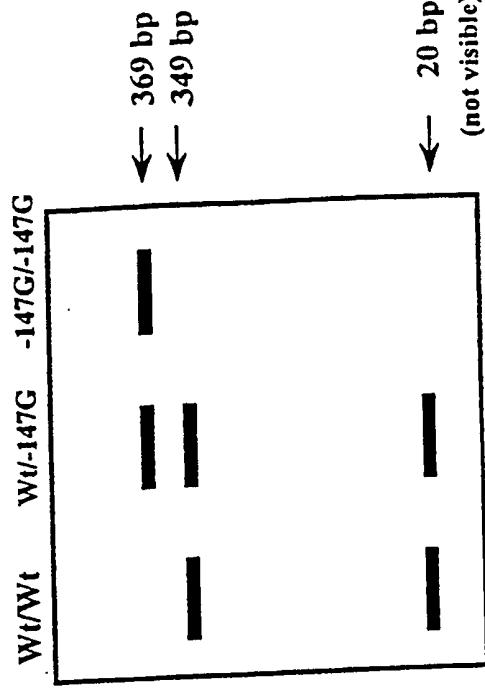


FIG. 3.

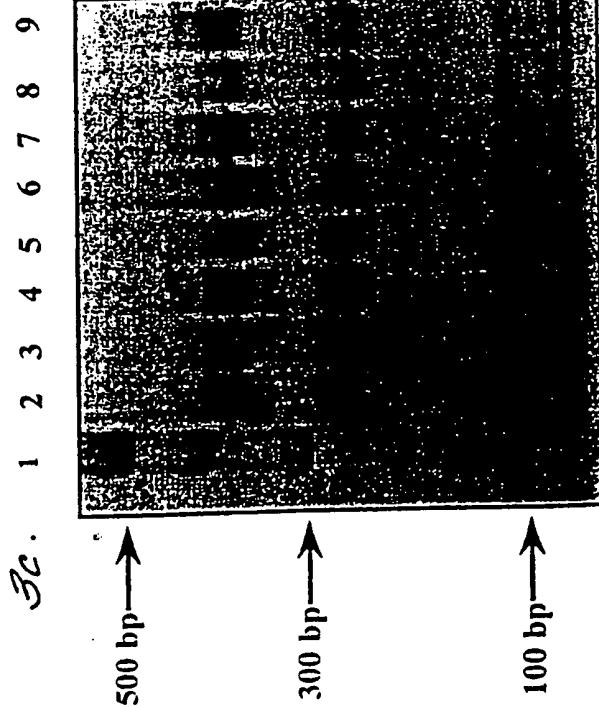
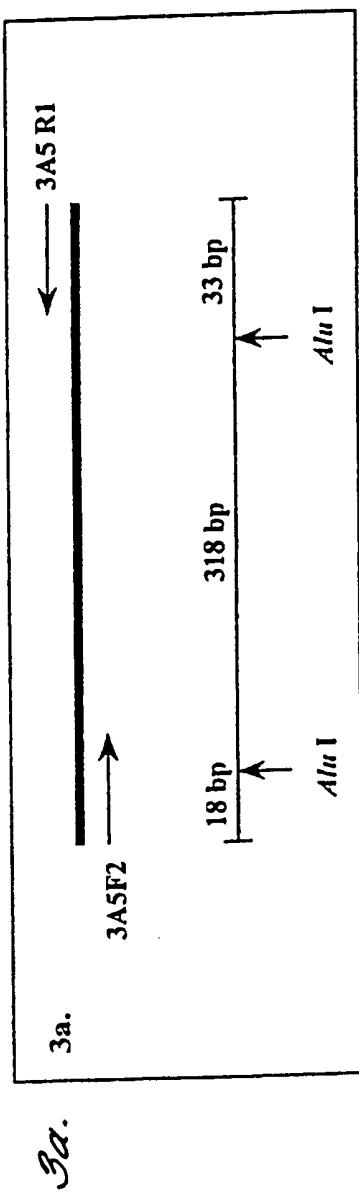


FIG. 4a.

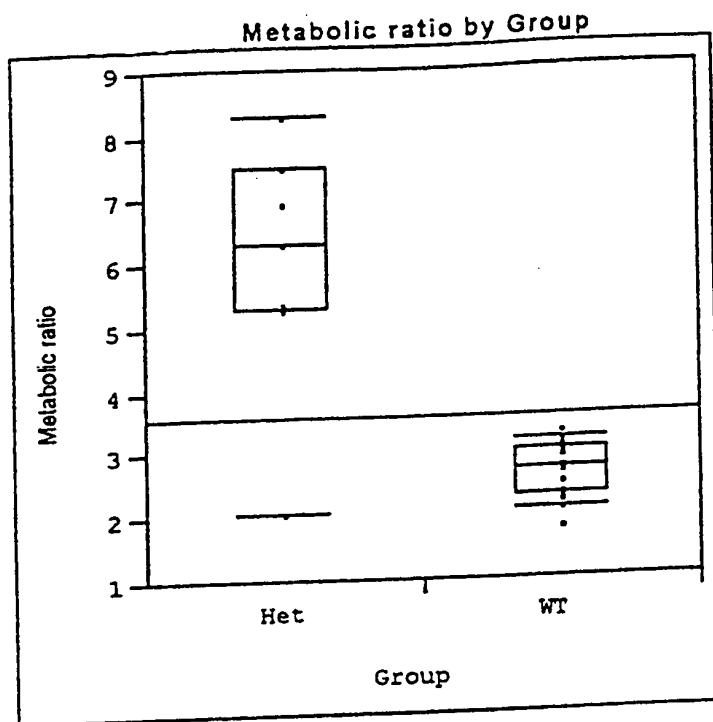


FIG. 4b.

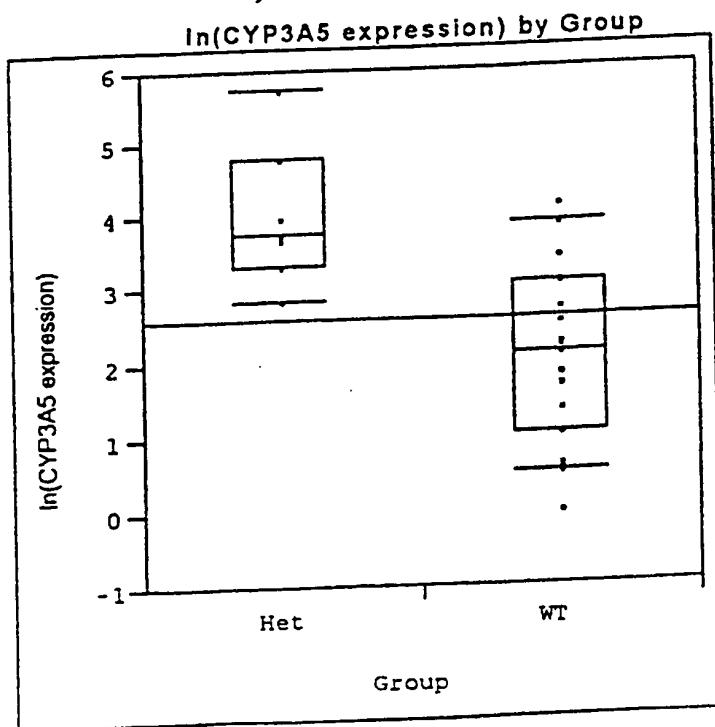
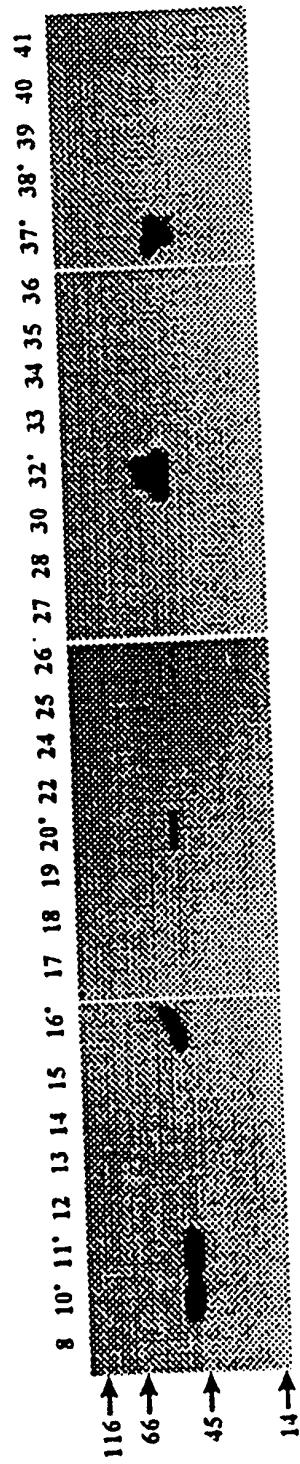


FIG. 5.



*FIG. 6.*

3A5F1 5'-GGGTCTGTCTGGCTGCGC-3'

3A5F2 5'-GGGGTCTGTCTGGCTGAGC-3'

3A5R1 5'TTTATGTGCTGGAGAAGGACG-3'

FIG. 7

-1343 GGAAGGCACC TACATGGCC TAAAACAGATG AATGGGTAAA GAGAGTACTT CACTTATGCA CAATGGGATA  
 -1273 CAATTCAGGC ATGAAAAAAG CATGAGATCC TGTCCTTTA ATAACGTGG CTGGAACTGC AGGTCATT  
 -1203 GTTAGGTTAA ATAGGCCAGG CACAAAGA CAGACATTC CATGGCTTCAC TTATTTGTGG GATCTACAAA  
 -1133 TCAAAACAT TGAGCTAATG TCTGGGTCTT AGTCAATTTT GTACCCTAAG TACGGGC ACAGCCATTA  
 -1063 GAATACATGA TGAATGTCTT ATACAGGAA TGAATAGGTGA AGGGCACG AGGGTTGG GGTTCTCTCG  
 -993 ATACATAGTA TCTTCTTGA CACATTCAGT ACAACTCTA ACAGGTTAAGT CTCTTCATGT ATGTTACCTT  
 -923 CTGAGGAAT AAGTTGGCAGA ACATGCCTTC TATATTTTCCTTTGGCAGA CAAGACCAAT TGCATTAGGTT  
 -853 GGGAAACGT GCTGGGTGCA TCTGAGGCCC AGCAAACATCA TAGTCTTATTGCAT CTATCACCAC AGACTCAGAG  
 -783 GGGATGAAC ACAGGGGGCCC AGCAATCTCACCCAAGTCA CTCCACCAAC ATTTCTGGTCG ACCACCCATG  
 -713 TGTACAGTAC CCTGTCTAGGG TCCAGGGGTCA TGAAGGTAA TAATACCAGA CTGTGCCCTT GAGGAACTCA  
 -643 CCTCTGTCAA GGGAAACAGG CACGAAACC CACAAGGGTG GTAGAGAGGA AATAGGACAA TAGGACTGTG  
 -573 TGAGGGGTAT AGGGGGCCACCA GAGGGAGGAAATGGTTTACAA TCTGTGTGAG GAGGTGGTA AGGGAAAGACT  
 -503 TTAATAGGAAG GGGTCTGTCT GGCTGGCTTT GCAGGAGATGT GTAGGTCA TCTAGGGGGCC ACAAGTACAC  
 -433 TCCAGGGCAGA GGGAATTGCA TGGTAAAGAAG TCTGCAGTTG GGATGGATTT CAAGTATCT  
 -363 GGAATGAGAAG CAGCCCATGGGA ACAAAGGGCA GGTGAGGGA TATTAAGAG GCTTCATGCCC ATGGGTTCCA  
 -293 CTTCAGGTTTC TGATAGAAAC TCAGGTTCCG TGGACTCCCT GATAAAACGTU ATAAGTTTGTTTATGATTCC  
 -223 CCATAGAATAAA TGAAACTCAAA GGAGGTAAGC AAGGGTAAGC GTGCAGGCTGG CTGGCAGGCTGG

## FIG. 7 (continued)

-153 AGCCCCACCT CCTTCTCCAG CACATARRACA TTTCAGCAGC TTGACCTAAG ACTGCTGTGC AGGGCAGGGA  
 -83 TGCTCCAGGC AGACAGGCCA GCAAACARACA GCACACAGCT GAAAGTAAGA CTCAGAGGAG ACAGTTGAAG  
 -13 AAGGCCAAGTG GCGATG

## Variant Sequences in the 5' flanking region of CYP3A5

Position	Variant sequence	Percentage
-1317	G-K (T or G) heterozygote	2.8% (1/36)
-988	T-Y (C or T) heterozygote	8.3% (3/36)
-657	C-Y (C or T) heterozygote	8.3% (3/36)
-475	T-K (T or G) heterozygote	30.6% (11/36)
-264	G-R (G or A) heterozygote	2.8% (1/36)
-147	A-R (A or G) heterozygote	30.6% (11/36)

FIG. 8.

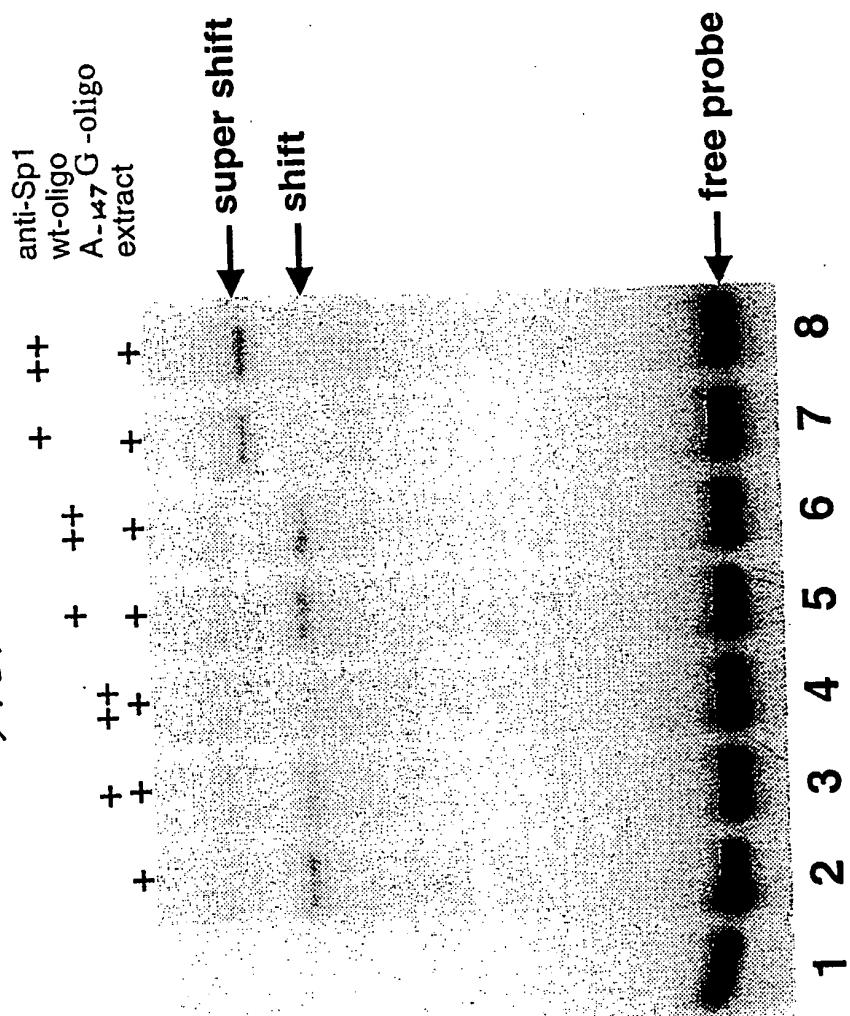


FIG. 9.

## CYP3A4, CYP3A5, CYP3A6/7

sites	* 20 * 40 * 60	
HSCYPFLA_CYP3A6/7	-----	
HSRCYP3_CYP3A7	-----	
HSP4503A4_CYP3A4	-----	
S74699_CYP3A5	-----	
S74700_CYP3A5	TAATTCGGTGTAATTTTAAATATAGAGAAGCTTGGGCCATTAATTC	60
sites	* 80 * 100 * 120	
HSCYPFLA_CYP3A6/7	-----	
HSRCYP3_CYP3A7	-----	
HSP4503A4_CYP3A4	-----	
S74699_CYP3A5	-----	
S74700_CYP3A5	ATTTTAAATTTTACCGATTTTAAAGATTTAGTAACTGGGATTTAATTC	120
sites	* 140 * 160 * 180	
HSCYPFLA_CYP3A6/7	-----	
HSRCYP3_CYP3A7	-----	
HSP4503A4_CYP3A4	-----	
S74699_CYP3A5	-----	
S74700_CYP3A5	GGGAGAACCGATAATATATTCAGCATACTTAACTGGT	180
sites	* 200 * 220 * 240	
----- [L1-RETROTRANSPOSON-ELEMENT] -----		
HSCYPFLA_CYP3A6/7	-----	
HSRCYP3_CYP3A7	-----	
HSP4503A4_CYP3A4	GATG-----GCTGAGGT--GGTTGGGTCATCTGGCTATC-----TGGGC	64
S74699_CYP3A5	GATTAACAAATGCCAAGATTGGAAGCAACCTACATGTCCATCAACAGATGAATGGGT	92
S74700_CYP3A5	GATTAACAAATGCCAAGATTGGAAGCAACCTACATGTCCATCAACAGATGAATGGGT	240
sites	* 260 * 280 * 300	
HSCYPFLA_CYP3A6/7	-----	
HSRCYP3_CYP3A7	-----	
HSP4503A4_CYP3A4	A-----GCTGTTCTCTT-----	76
S74699_CYP3A5	AAAGAGAGTACTTCACTTATGCACAATGGAGTACAATTCAAGCCATGAAAAAAGCATGAGA	152
S74700_CYP3A5	AAAGAGAGTACTTCACTTATGCACAATGGAGTACAATTCAAGCCATGAAAAAAGCATGAGA	300
sites	* 320 * 340 * 360	
HSCYPFLA_CYP3A6/7	-----	
HSRCYP3_CYP3A7	-----	
HSP4503A4_CYP3A4	--CTCTCCCTTTCT-----CTCCCTGTT	96
S74699_CYP3A5	TCCTGTCCCTTTATAATAACGTGGCTGAACTCAGGTCATTATGTTAGGTAAAATAAG	212
S74700_CYP3A5	TCCTGTCCCTTTATAATAG---CGTGGCTGGACTGCAGGTCATTATGTTAGGTAAAATAAG	357
sites	* 380 * 400 * 420	
----- [L1-RETROTRANSPOSON-ELEMENT] -----		
HSCYPFLA_CYP3A6/7	-----	
HSRCYP3_CYP3A7	-----	
HSP4503A4_CYP3A4	CCAGACATGCAG-----TATTT-----	118
S74699_CYP3A5	CCAGGCACACAAAGACAGACATTGCATGTTCTCACTTATTGTTGGGATCTAATTC	272
S74700_CYP3A5	CCAGGCACACAAAGACAGACATTGCATGTTCTCACTTATTGTTGGGATCTAATTC	417

FIG. 9 (CONTINUED 1).

## CYP3A4, CYP3A5, CYP3A6/7

sites							
HSCYPFLA_CYP3A6/7	:	440	*	460	*	480	:
HSRCYP3_CYP3A7	:						168
HSP4503A4_CYP3A4	:	GGATGGGGATTTTTCGAAAGTATGTCTTGAGTCGGT					332
S74699_CYP3A5	:	ATTTTACGATATGTCTGGGTATTTGTAAATTTGTA					477
S74700_CYP3A5	:	ATTTTACGATATGTCTGGGTATTTGTAAATTTGTA					
sites							
HSCYPFLA_CYP3A6/7	:	500	*	520	*	540	:
HSRCYP3_CYP3A7	:						207
HSP4503A4_CYP3A4	:	ATCCGTTTTCGCGATTTGAGGATGCGATGCGATG					391
S74699_CYP3A5	:	ATTTGATTAATCTTGATGATGCTTAAACGCTTAC					537
S74700_CYP3A5	:	ATTTGATTAATCTTGATGATGCTTAAACGCTTAC					
sites							
HSCYPFLA_CYP3A6/7	:	560	*	580	*	600	:
HSRCYP3_CYP3A7	:						255
HSP4503A4_CYP3A4	:	ATATTTATTAATGATGATGATGATGATGATGATG					450
S74699_CYP3A5	:	GTTCGGGTCGTTTGTGTTATGTTGTTGTTGTTG					597
S74700_CYP3A5	:	GTTCGGGTCGTTTGTGTTATGTTGTTGTTGTTG					
sites							
HSCYPFLA_CYP3A6/7	:	620	*	640	*	660	:
HSRCYP3_CYP3A7	:						303
HSP4503A4_CYP3A4	:	TTTTGGGCTTTAGAAAGATTTGCTTGGCTTGGCT					508
S74699_CYP3A5	:	GCTTAAGTTTTCAGTTGATGTTAATAGGCTTAA					656
S74700_CYP3A5	:	G-TAGGTATTTAGTATGATGATGATGATGATG					
sites							
HSCYPFLA_CYP3A6/7	:	680	*	700	*	720	:
HSRCYP3_CYP3A7	:						349
HSP4503A4_CYP3A4	:	CGAGGAGAGTGTCTGGCTTACGCTGGCTTGGCT					568
S74699_CYP3A5	:	TTTATTTATTTAGGAGAGAAATTGCTTGGGAA					716
S74700_CYP3A5	:	TTTATTTATTTAGGAGAGAAATTGCTTGGGAA					
sites							
HSCYPFLA_CYP3A6/7	:	740	*	760	*	780	:
HSRCYP3_CYP3A7	:						391
HSP4503A4_CYP3A4	:	TGAGCTTACGCTTACGCTTACGCTTACGCTTACG					628
S74699_CYP3A5	:	ATGCTTACGCTTACGCTTACGCTTACGCTTACG					776
S74700_CYP3A5	:	ATGCTTACGCTTACGCTTACGCTTACGCTTACG					
sites							
HSCYPFLA_CYP3A6/7	:	800	*	820	*	840	:
HSRCYP3_CYP3A7	:						451
HSP4503A4_CYP3A4	:	GGGGCTTACGCTTACGCTTACGCTTACGCTTACG					688
S74699_CYP3A5	:	AGGGCTTACGCTTACGCTTACGCTTACGCTTACG					836
S74700_CYP3A5	:	AGGGCTTACGCTTACGCTTACGCTTACGCTTACG					

FIG. 9 (CONTINUED 2)

### CYP3A4, CYP3A5, CYP3A6/7

sites HSCYPFLA\_CYP3A6/7 : \* 860 \* 880 \* 900 : -  
 HSRCYP3\_CYP3A7 : -  
 HSP4503A4\_CYP3A4 : -  
 S74699\_CYP3A5 : -  
 S74700\_CYP3A5 : -

sites HSCYPFLA\_CYP3A6/7 : \* 920 \* 940 \* 960 : -  
 HSRCYP3\_CYP3A7 : -  
 HSP4503A4\_CYP3A4 : -  
 S74699\_CYP3A5 : -  
 S74700\_CYP3A5 : -

sites HSCYPFLA\_CYP3A6/7 : \* 980 \* 1000 \* 1020 : -  
 HSRCYP3\_CYP3A7 : -  
 HSP4503A4\_CYP3A4 : -  
 S74699\_CYP3A5 : -  
 S74700\_CYP3A5 : -

sites HSCYPFLA\_CYP3A6/7 : \* 1040 \* 1060 \* 1080 : -  
 HSRCYP3\_CYP3A7 : -  
 HSP4503A4\_CYP3A4 : -  
 S74699\_CYP3A5 : -  
 S74700\_CYP3A5 : -

sites HSCYPFLA\_CYP3A6/7 : \* 1100 \* 1120 \* 1140 : -  
 HSRCYP3\_CYP3A7 : -  
 HSP4503A4\_CYP3A4 : -  
 S74699\_CYP3A5 : -  
 S74700\_CYP3A5 : -

sites HSCYPFLA\_CYP3A6/7 : \* 1160 \* 1180 \* 1200 : -  
 HSRCYP3\_CYP3A7 : -  
 HSP4503A4\_CYP3A4 : -  
 S74699\_CYP3A5 : -  
 S74700\_CYP3A5 : -

sites HSCYPFLA\_CYP3A6/7 : \* 1220 \* 1240 \* 1260 : -  
 HSRCYP3\_CYP3A7 : -  
 HSP4503A4\_CYP3A4 : -  
 S74699\_CYP3A5 : -  
 S74700\_CYP3A5 : -

## FIG. 9 (CONTINUED 3).

## CYP3A4, CYP3A5, CYP3A6/7

sites	* 1280	* 1300	* 1320	
HSCYPFLA_CYP3A6/7	[REDACTED]			
HSRCYP3_CYP3A7	[REDACTED]			916
HSP4503A4_CYP3A4	[REDACTED]			1165
S74699_CYP3A5	[REDACTED]			1307
S74700_CYP3A5	[REDACTED]			
sites	* 1340	* 1360	* 1380	
HSCYPFLA_CYP3A6/7	[REDACTED]			
HSRCYP3_CYP3A7	[REDACTED]			976
HSP4503A4_CYP3A4	[REDACTED]			1224
S74699_CYP3A5	[REDACTED]			1366
S74700_CYP3A5	[REDACTED]			
sites	* 1400	* 1420	* 1440	
HSCYPFLA_CYP3A6/7	[REDACTED]			
HSRCYP3_CYP3A7	[REDACTED]			55
HSP4503A4_CYP3A4	[REDACTED]			1036
S74699_CYP3A5	[REDACTED]			1242
S74700_CYP3A5	[REDACTED]			1383
sites	* 1460	* 1480	* 1500	
HSCYPFLA_CYP3A6/7	[REDACTED]			
HSRCYP3_CYP3A7	[REDACTED]			62
HSP4503A4_CYP3A4	[REDACTED]			115
S74699_CYP3A5	[REDACTED]			1096
S74700_CYP3A5	[REDACTED]			1286
sites	* 1520	* 1540	* 1560	
HSCYPFLA_CYP3A6/7	[REDACTED]			40
HSRCYP3_CYP3A7	[REDACTED]			174
HSP4503A4_CYP3A4	[REDACTED]			1156
S74699_CYP3A5	[REDACTED]			1345
S74700_CYP3A5	[REDACTED]			1486
sites	* 1580	* 1600	* START-ATG	
HSCYPFLA_CYP3A6/7	[REDACTED]			70
HSRCYP3_CYP3A7	[REDACTED]			94
HSP4503A4_CYP3A4	[REDACTED]			228
S74699_CYP3A5	[REDACTED]			1210
S74700_CYP3A5	[REDACTED]			1399
	[REDACTED]			1540
	ac	gctaaaaa	aagactcagaggaga	ag t a aagg aaGT G gATG

*FIG. 9(A).*

A or G  
 =  
 MEME repeated motif 9  
 =====  
 MEME 'single' motif 9  
 =====  
 Yi-consensus  
 =====  
 apoE-undefined-site-3  
 =====  
 ApoE\_B1  
 =====  
 APRT-human\_US  
 =====  
 APRT-CHO\_US  
 =====  
 1238 AGCTGCAGCCCCACCTCCTTCTCCAGC  
 TCGACGTCGGGGTGGAGGAAGAGTCG

(B) A

*FIG. 9(B).*

MEME repeated motif 2  
 =====  
 MEME repeated motif 2  
 =====  
 MEME 'single' motif 9  
 =====  
 Yi-consensus  
 =====  
 Sp1-TPI\_(4)  
 =====  
 GCF-consensus  
 =====  
 DSE\_(1)  
 =====  
 IRE\_(1)  
 =====  
 Sp1\_CS4  
 =====  
 GC-box\_(1)  
 =====  
 Sp1-IE-4/5  
 =====  
 Sp1-IE-3.3  
 =====  
 E2A\_CS hsp70.2  
 =====  
 E2A\_CSSp1-hsp70\_(1)  
 =====  
 APRT-mouse\_US  
 =====  
 1379 AGCTGCAGCCCCGCCTCCTTCTCCAGC  
 TCGACGTCGGGGCGGAGGAAGAGTCG

(B) G

FIG. 9(c).

MEME repeated motif 9  
 =====  
 MEME repeated motif 9  
 =====  
 MEME repeated motif 3  
 =====  
 MEME 'single' motif 6  
 =====  
 E-2.7\_kb\_(3)  
 =  
 E1A-F\_CS  
 =====  
 GH1  
 =====  
 MTVGRE\_NRS  
 =====  
 910 TCTGCTGGCTGGCTTGCAAGGATGTGTAG  
 AGACAGACCGACCCGAAACGTTCTACACATC  
 .....

(A)<sub>T</sub>

FIG. 9(d).

MEME repeated motif 9  
 =====  
 MEME repeated motif 9  
 =====  
 MEME repeated motif 9  
 =====  
 MEME repeated motif 3  
 =====  
 MEME 'single' motif 6  
 =====  
 E-2.7\_kb\_(3)  
 =  
 MBF-I\_CS  
 =====  
 E1A-F\_CS  
 =====  
 CNBP-SREMVGRE\_NRS  
 =====  
 GH1 MRE\_CS2  
 =====  
 1052 TCTGCTGGCTGGCTGCAAGGATGTGTAG  
 AGACAGACCGACCCGACGTTCTACACATC  
 .....

(A)<sub>G</sub>